Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research
Notice Number: NOT-OD-23-091

Name: Chris Bourg
Name of Organization: Massachusetts Institute of Technology - MIT Libraries
Type of Organization *: University
Role: Institutional Official

NIH seeks information regarding the NIH Public Access Plan, from all interested individuals and communities, including, but not limited to, authors, investigators, research institutions, libraries, scholarly publishers, scientific societies, healthcare providers, patients, students, educators, research participants, and other members of the public. While comments are welcome on all elements of the NIH Public Access Plan, input would be most welcome on Section III related to scholarly publications and on the particular issues identified below. Comments may be entered below or attached in the next section.

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

- The current incentive structures in science, higher education, and their influence on scholarly publishing more broadly, continue to be the major source of the inequities that manifest in publication opportunities for NIH-supported investigators. Across the globe there are groups working to address this important issue such as the Higher Education Leadership Initiative for Open Scholarship (HELIOS), the Declaration on Research Assessment (DORA), GRC research assessment working group, The Latin American Forum on Research Assessment (FOLEC-CLACSO), the UKRI Future research assessment program, the Dutch initiative “Room for everyone’s talent”, HuMetrics HSS, the Research Data Alliance (RDA) Evaluation of Research Interest Group, and multiple initiatives under the European Research Area policy agenda, as reforming research assessment is one of their 20 priority actions. The current inequity in the scholarly communications system, either in the subscription model, which blocks access to publications, or the APC model, which blocks access to publishing, is driven by incentive structures that reward the “publish or perish” mentality and reduce the assessment of research to a list of publications valued according to the journal they are published in. This sidelines the assessment of both the quality of the research itself and the actual

components that are created as part of the research process. Therefore, the NIH should amplify the importance of these efforts and seek to encourage the implementation of equitable research assessment practices and career advancement incentives that will increase equity in science and in publication opportunities.

- The financial burden placed on investigators by several scholarly publication models is also an important source of inequity in publication opportunities. In most cases, the Article Processing Charge (APC) model has proven prohibitively expensive while providing very poor value for money and diverting funds that could be better applied towards more critical research needs. As Grossman and Brembs (2021) highlight, these costs often include the value of the investigator labor and research already included in their submission and the volunteer labor of editors and peer reviewers. To the extent that can be determined given the lack of cost transparency, the actual publisher value-added services make up a very small portion of the overall costs included in that charge. As stewards of taxpayer monies, the NIH should avoid being charged twice by hybrid journals.

  To ensure equity in this area, it is important that the NIH make it clear that there are cost-free paths towards compliance, and that neither researchers nor institutions should feel compelled to purchase their way towards compliance. They should also provide their investigators with language and specific guidance to help ensure that authors retain the rights necessary to make their federally funded, peer-review manuscript freely available and reusable post-publication in PMC – without an embargo period – to be in compliance with the NIH's policy.

- Given the inequities inherent in the pay-to-publish model prevalent among the majority of publishers of federally funded research, a shift towards greater reliance on the well-established network of repositories is critical to an equitable implementation of NIH's public access plan. (see U.S. Repository Network's Desirable Characteristics of Digital Publication Repositories and COAR Community Framework for Good Practices in Repositories). The repository path eliminates author-facing financial burdens and reduces the inequities in publication opportunities. In addition, repository services typically include a level of expertise in the curation, discovery, and reusability of scholarly content that is lacking across the commercial publishing landscape.

  MIT's institutional repository has enabled 58% of all faculty publications since 2009 to be publicly available. This has been achieved only through MIT resourcing the libraries robustly enough to support monitoring, support, and outreach to faculty about the Institute's policy, and equally resourcing the repository infrastructure and the technical, metadata, and expertise needed. The cost to sustainably support public access to NIH-funded research would be significantly less if it were implemented through repositories compared to the costs of the current commercial scholarly publication models which funnels significant amounts of taxpayer dollars into commercial publishers' profit margins. Supporting repositories in this function would keep the resources within the

research system and would provide the opportunity for more equity in publication opportunities, as repositories do not usually charge authors to deposit their peer-reviewed manuscript.

2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

- Reproducibility is an equity issue. All of the necessary components to reproduce an experiment or a study need to be accessible in meaningful ways in order to ensure equity of opportunity to contribute in a field. Equitable, immediate access to a publication is a good but insufficient step to achieve the NIH's and the OSTP stated goals. Resources should be allocated for infrastructural support at a systematic level to be able to communicate, discover, and maintain the individual components of research in appropriate ways for those components.

- CC BY or similar licenses should be used to ensure that legal access for adaptation for accessibility concerns is permitted from the beginning. This kind of licensing would also permit the content to be translated into other languages which increases the potential audience and impact of NIH-funded research.

- While we applaud NIH's commitment to improving guidance for submitters on supplying more human-accessible content, it is important to note that such guidance should be sufficient to ensure that the full research product (text, figures, tables, scientific notations, etc.) is accessible to minimize the dependence on consumer accessibility remediation that may be difficult or limiting due to, for example, missing contextualization. As Brinn et al. (2022) note, publication accessibility too often falls short, and the NIH's guidance needs to be comprehensive and reflective of current and evolving accessibility approaches for publications.

- As mentioned above, ensuring that NIH investigators keep the rights necessary to make their final, peer-reviewed manuscript freely available and fully reusable post-publication in PMC without an embargo period is a critical step to achieve the NIH's stated goal. At the same time, it is important to require licenses that permit computational access of these publications for further research purposes.

3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

Listed publication fees, publish & subscribe agreements, and fee waivers have a substantial influence on publication opportunities in science. Evidence suggests that opportunities to publish systematically vary by race, gender of author, and characteristics of home institution. Measurement and evaluation, however, is currently obstructed by a lack of systematic open information about publication fees and author characteristics[1]

NIH can play a central role in addressing these gaps through specific, practical, and systematic actions. These include:

- requiring that NIH-supported publications include standard metadata documenting the standard/list APC used for that paper, the actual APC charged, and, if different, the amount of difference due to individual waiver and/or institutional agreement (e.g., publish and subscribe); and
- integrating demographic information NIH collects on awardees with data collected on publications to produce systematic public statistics on the distribution of publication fees for NIH outputs.

NIH efforts should be inclusive of research output beyond publication to include data management, sharing, and curation costs. To this end, we recommend that NIH:

- support research into data sharing costs – potentially utilizing data acquired from NIHDMSP – to increase understanding and estimates of costs and to determine factors that create equity hardships;
- require archives to report charges for publicly archived data (parallel to reporting APCS);
- analyze budgets from funded grants to inform understanding of data-sharing cost trends; and
- refer to those disciplines that have been sharing data effectively and efficiently for many years, and adopt and adapt the practices that are working in those disciplines.

4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

[1] See for a discussion Altman, Micah. "Designing Community Tracking Indicators for Open and Inclusive Scholarship." *Proceedings of the Association for Information Science and Technology* 59, no. 1 (2022): 393-397.

There is broad consensus in the scientific community that implementations of PIDs and metadata describing research inputs and outputs should be sufficiently robust to large scale (machine analysis, accountability and, where applicable, reuse of the referenced content).[2] Best practices in this area requires that selected PIDs' protocols and metadata formats are community-based and openly documented; each PID can be persistently, publicly, globally resolved to machine-actionable metadata; and the accompanying metadata provides sufficient information to enable direct machine and human access to the content of the described outputs (for authorized users).

To be consistent with these principles and practices, all NIH awards, awardees, and outputs should be associated with PIDs and metadata. Thus, NIH awards and any outputs from them (including preprints, publications, data, and software) should be associated with PIDs and metadata sufficient to: (a) locate its content and determine its accessibility; (b) link each output to any supporting NIH awards (and vice-versa); (c) link each individual NIH-supported publication to preprint versions, supporting data, and supporting software.

For this purpose we recommend that NIH adopt practices already used in the community for identifying and citing scientific publications, datasets and software. NIH should consider specific application of DOI's or the use of RAiDs (Research Activity Identifiers) for its awards and the research activities therein. These identifiers should resolve to standardized machine-actionable metadata, as per Starr et al. (2015). Research outputs referenced in awards' reporting should be associated with Funder Registry metadata and ORCID identifiers documenting all contributors, and include citation metadata. These outputs should include relevant data, research software, unique script or code, and other materials necessary to understand, validate, and support the research findings associated with the award.

[2] See for applications to data and software; Altman, Micah, Christine Borgman, Mercè Crosas, and Maryann Matone. "An introduction to the joint principles for data citation." *Bulletin of the Association for Information Science and Technology* 41, no. 3 (2015): 43-45.; Smith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group. (2016) Software Citation Principles. PeerJ Computer Science 2:e86.DOI: 10.7717/peerj-cs.86 ; Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific data* 3, no. 1 (2016): 1-9.